Proceedings of the Iowa Academy of Science

Volume 54 | Annual Issue Article 39

1947

Selection of the Optimal Scoring Plan for the Fritz Test of Cynicism

Charles O. Neidt
Iowa State College

Selection of the Optimal Scoring Plan for the Fritz Test of Cynicism*

CHARLES O. NEIDT

In an investigation involving a measuring instrument made up of situations to which a subject reacts on a response scale, consideration must be given the arithmetic weight-values to be assigned to the various responses along the scale before scoring the test. The present investigation was undertaken to select a scoring plan, i.e., weight-values, for the response scale provided in the Fritz Test of Cynicism.

Preliminary Investigation

To investigate the nature of the cynical attitudes held by college students, 200 test items were originated by Dr. M. F. Fritz of the Department of Psychology and the Testing Bureau of Iowa State College. The items were stated in an informal style and difficult vocabulary words were either explained or avoided.

It was assumed by Fritz that past experience of the subjects would result in an attitude of acceptance or rejection toward each test item, and that this attitude could be expressed quantitatively, that is, mildly or strongly. Since Fritz considered the "intermediate" or "undecided" attitude as rarely existent, and to force the subject to reflect as nearly as possible the direction of his reaction to each test item, a four-point response scale was provided for the subject to indicate the strength of his acceptance or rejection of each of the test items.

The test items were so arranged that a cynical response for some items would be indicated by agreeing with the item, and for the others by disagreeing with the item. This procedure is illustrated by the following sample items:

"I would say that perhaps as much as half of our tax money finds its way into the hands of grafters.....A a d D

Response Scale

A—*Definitely acceptable. I agree most wholeheartedly with the statement.*

a—*Fairly acceptable. In general, I agree with the statement ment but not too strongly.*

d—*Mild disagreement. I am inclined to agree with the statement. In general, I think I would reject it.*

D—*Definitely rejected. I absolutely do not agree with the statement.*

*This report is part of a study for which a grant-in-aid was received from the Iowa Academy of Science.

"I believe that at least 90% of the girls would rather marry a poor boy whom they love than a rich man whom they do not love" ..... A a d D

Acceptance of the first test item and rejection of the second were considered to have the same interpretation regarding cynicism.

WEBSTER'S COLLEGIATE DICTIONARY defines cynical as "contemptuously distrustful of human nature" and a cynic as "one who believes that human conduct is motivated wholly by self-interest". The two cynical responses for each of the 200 test items were validated in terms of these definitions.

Copies of the test were distributed at random intervals to 400 students enrolled in psychology courses at Iowa State College during the year 1944 and 1945. Students were requested to complete a short questionnaire attached to the test regarding their sex, political preference, church preference, father's occupation, marital status, education, and age. The tests were completed and returned at the convenience of the students.

To facilitate scoring the tests, the responses of the subjects were transferred to electrically scored answer sheets, and all scoring was accomplished with the use of an International Business Machine Test Scorer.

Criteria of Selection Involved

In selecting the best scoring plan for an attitude scale, an external criterion, or the extent to which the plan will result in total scores which will differentiate between groups of subjects possessing known differences of the characteristic measured by the scale, is of great importance. It seems reasonable to expect the optimal scoring plan to differentiate to the greatest extent between the groups when scoring plans are compared. This "differentiating ability" can be determined by scoring each scale by the several plans under consideration and applying a test of significance to the total scores of the differing groups.

In many testing investigations groups of subjects possessing the desired differences of the characteristic being measured can be identified relatively easily, whereas differing groups in other investigations may not lend themselves to identification in this manner. Under ideal conditions the differing groups are identifiable before or during the administration of the test. If, however, very little is known about the characteristic being measured by the test, it may be necessary to employ a semi-external criterion, that is, to identify the differing groups after the test scores have been examined. In the investigation of cynicism, groups possessing known differences of the characteristic were not available before the administration of the test and a semi-external criterion was therefore used.

The results of a preliminary study reported by the author at the 1946 session of the Iowa Academy of Science, indicated the existence

of highly significant differences between sexes and among ages for the scores of 364 subjects. The test for significance revealed greater differences between sexes than among ages, although both tests were significant at the one per cent level of confidence. On the basis of the preliminary study it was assumed that men reflect more cynicism in their attitudes than women, and that older students reflect more cynicism than younger students.

For the present investigation it was assumed that the scoring plan which would yield the greatest differences between sexes and among ages on the basis of total scores would provide evidence for the selection of the optimal scoring plan for the Fritz Test of Cynicism.

The four-point response scale used with the test of cynicism made possible several weight-value combinations, and from these the four scoring plans which were considered feasible were chosen. Assigning weight-values of 2, 1, 0, 0 to the response scale—the highest value being assigned to the most cynical of the four responses—was designated as the first scoring plan; 3, 2, 1, 0 as the second; 4, 3, 1, 0 as the third; and 1, 0, 0, 0 as the fourth.

The 400 tests were scored and weighted by each of the four scoring plans and the total scores classified by sex and age of the subject. The average score and frequency classified by sex and age for the four scoring plans are shown in Tables 1, 3, 5 and 7. Data were not available concerning the age of 13 subjects. The scores of these subjects were disregarded for purposes of computation.

The scores obtained by each of the scoring plans were treated by analysis of variance to test the significance of the differences revealed. The data were treated for disproportionality following the method suggested by Snedecor (1). The analysis of variance tables and F-values for the four scoring plans are shown in Tables 2, 4, 6 and 8.

Table 1

Average Score by the 2-1-0-0 Scoring Plan for the Various Ages Classified by Sex

	Male		Female	
Age	N	Average	N	Average
17	13	93.077	18	80.667
18	19	93.895	73	80.575
19	10	94.500	82	69.890
20	14	92.857	42	68.000
21	13	96.846	26	88.154
22	11	110.273	7	83.286
23	7	103.857	3	63.667
24	14	94.857	5	97.600
25 +	25	86.230	5	86.800
Total	126	96.266	261	79.849

Table 2

Analysis of Variance of 2,1,0,0 Scoring Plan

Source	d.f.	Sum of Squares		Mean Square	F
		Unadjusted	Adjusted		
Sex	1	28,630.00	16,875.51	16,875.51	28.87**
Age	8	26,110.14	14,355.65	1,794.46	3.07**
Sex x Age	8	128,019.66	6,484.85	810.61	1.39
Unexplained	369	215,708.20	215,708.20	584.58	
Total	386	398,468.00			

**Significant at the one per cent level of confidence.

Table 3

Average Score by the 3,2,1,0 Scoring Plan for the Various Ages Classified by Sex

Age	Male		Female	
	N	Average	N	Average
17	13	233.615	18	202.389
18	19	219.158	73	203.589
19	10	227.900	82	185.146
20	14	222.786	42	188.000
21	13	240.462	26	211.192
22	11	229.545	7	203.143
23	7	248.286	3	166.334
24	14	220.071	5	234.400
25 +	25	207.640	5	203.000
Total	126	224.286	261	196.100

Table 4

Analysis of Variance of 3,2,1,0 Scoring Plan

Source	d.f.	Sum of Squares		Mean Square	F
		Unadjusted	Adjusted		
Sex	1	67,510.20	45,487.47	45,487.37	16.75**
Age	8	52,747.00	30,724.17	3,840.52	1.41
Sex x Age	8	-2,318.60	20,394.30	2,549.29	.94
Unexplained	369	1,002,029.40	1,002,029.40	2,715.53	
Total	386	1,119,968.00			

**Significant at the one per cent level of confidence.

Table 5

Average Score by the 4,3,1,0 Scoring Plan for the Various Ages Classified by Sex

	Male		Female	
Age	N	Average	N	Average
17	13	303.769	18	261.278
18	19	287.000	73	232.241
19	10	295.000	82	239.183
20	14	287.786	42	246.619
21	13	318.000	26	274.423
22	11	301.545	7	264.286
23	7	321.429	3	214.000
24	14	286.071	5	303.600
25 +	25	275.960	5	268.400
Total	126	293.540	261	216.234

Table 6

Analysis of Variance of 4,3,1,0 Scoring Plan

Source	d.f.	Sum of Squares		Mean Square	F
		Unadjusted	Adjusted		
Sex	1	129,196.00	82,300.31	82,300.31	16.09**
Age	8	97,424.10	50,528.41	6,316.05	1.24
Sex x Age	8	-14,063.50	32,189.50	4,023.69	.79
Unexplained	369	1,886,988.40	1,886,988.40	5,113.79	
Total	386	2,099,545.00			

**Significant at the one per cent level of confidence.

Table 7

Average Score by the 1,0,0,0 Scoring Plan for the Various Ages Classified by Sex

	Male		Female	
Age	N	Average	N	Average
17	13	22.923	18	21.778
18	19	26.053	73	20.384
19	10	26.400	82	17.073
20	14	26.786	42	14.381
21	13	23.308	26	26.346
22	11	38.455	1	22.143
23	7	30.714	3	16.000
24	14	29.000	5	28.400
25 +	25	21.680	5	21.400
Total	126	26.357	261	19.238

Table 8

Analysis of Variance of 1,0,0,0 Scoring Plan

Source	d.f.	Sum of Squares		Mean Square	F
		Unadjusted	Adjusted		
Sex	1	4,307.36	2,022.01	2,022.01	9.06**
Age	8	7,992.40	5,970.39	746.30	3.34**
Sex x Age	8	212.79	2.274.71	284.34	1.27
Unexplained	369	82,383.01	82,383.01	223.26	
Total	386	94,895.56			

**Significant at the one per cent level of confidence.

A summary of the F-values for the four scoring plans is shown in Table 9. Inspection of this table indicates that the scoring plan of 2, 1, 0, 0 differentiates to the greatest extent between sexes, and the scoring plan of 1, 0, 0, 0 differentiates to the greatest extent among ages, but only to a slightly greater extent than the 2, 1, 0, 0 method.

Table 9

F-values for Sex, Age, and Sex x Age for Four Scoring Plans

Scoring Plan	F-value		
	Sex	Age	Sex x Age
2,1,0,0	28.87	3.07	1.39
3,2,1,0	16.75	1.41	.94
4,3,1,0	16.09	1.24	.79
1,0,0,0	9.08	3.34	1.27
Required for 1%	6.70	2.55	2.55
Significance 5%	3.86	1.96	1.96

In comparing the scoring plans, consideration must be given to the possibility that individuals in one group may show a tendency to mark more extreme responses possibly due to less cautiousness and hence raise their scores because of this tendency rather than because of actual possession of a larger amount of the characteristic measured. Obviously, if both ends of the response scale provided for each test item are used in each scoring plan under consideration, the bias will tend to be eliminated.

The 3, 2, 1, 0 and the 4, 3, 1, 0 scoring plans use both ends of the response scale and the 2, 1, 0, 0 and the 1, 0, 0, 0 scoring plans do not. To determine whether either sex showed a tendency to mark more extremes, weight-values of arithmetic sequence were assigned to points along the scale for each sex according to the number of responses made at each point of the response scale as shown in Table 10. The standard deviations were computed for each sex using the responses of the 400 subjects and found to be 1.0351 for males and

Table 10

Responses to the Four-Point Scale by Sex

Sex	N	Type of Response: Strongly Cynical	Mildly Cynical	Mildly Opposite	Strongly Opposite	Total
Male	130	3,414	5,375	8,022	9,022	25,841
Female	170	5,160	10,246	17,346	20,925	53,677
Total	400	8,584	15,619	25,368	29,947	79,518

.98244 for females. The F-value for significance of the difference between these two standard deviations was found to be 1.0035, with 25,840 and 53,676 degrees of freedom. Since the numbers of degrees of freedom were as large as indicated, this value was found to be highly significant. It was concluded that the difference between the sex F-value of the 2, 1, 0, 0 scoring plan and the 3, 2, 1, 0 plan was partially the result of a tendency of men to mark more extreme responses than women. No attempt was made to determine the proportion of the difference resulting from this tendency of the men to mark extremes, nor were the standard deviations for the various age groups computed. Since the proportion of the sex F-value of the 2, 1, 0, 0 scoring plan resulting from the tendency of men to mark extreme responses was undetermined, it was concluded that the difference between the F-value for sex of the 2, 1, 0, 0 scoring plan and the F-value for sex of the 3, 2, 1, 0 scoring plan was great enough to justify the 2, 1, 0, 0 scoring plan regarding differential ability between sexes.

To test the possibility that one scoring plan might reveal a higher coefficient of reliability than the other plans, a proportionate sample of 17 miles and 33 females was selected from the 387 cases under consideration by using a table of random numbers. The odd-even coefficients of reliability were computed for each of the four scoring plans. The coefficients of reliability for the entire test were estimated by the Spearman-Brown modified formula and the results are shown in Table 11.

Table 11

Coefficients of Reliability for Four Scoring Plans

Scoring Plan	Odd-Even r	Estimated r
2,1,0,0	.8273	.9054
3,2,1,0	.8433	.9150
4,3,1,0	.8098	.8949
1,0,0,0	.8365	.9110

Inspection of this table indicates that the differences among the coefficients of reliability for the four scoring plans are negligible. It was concluded that each method of scoring yielded a high co-

efficient of reliability, but that the differences among the four coefficients were so small that no distinct scoring advantage concerning reliability was revealed for any one scoring plan.

To determine whether or not the weight-values for the responses should follow an arithmetic sequence with a difference of one between each weight-value, and to determine internally consistent weight-values for the responses, a normal distribution of responses for the group of subjects was assumed and the sigma scores of percentages were computed for each response as shown in Table 12.

Examination of these data indicates that if the cynicism measured by the test follows a normal distribution for this group of subjects, the weight-values based on sigma scores of percentages closely ap-

Table 12

Weights for Responses—Normal Curve

	Cynical Strongly	Cynical Mildly	Opposite Mildly	Opposite Strongly	Total
Total Responses	8,584	15,619	25,368	29,947	79,518
% of Total	10.8%	19.6%	31.9%	37.7%	
Area Values	94.6%	74.9%	53.65%	18.8%	
Sigma Values	1.6072	.8204	.0918	-.8853	
Adjusted Sigma Values	2.4925	1.7057	.9771	0	
Adjusted Scale Values from 0 to 3	3.00	2.06	1.17	0	

proximate an arithmetic sequence of 0, 1, 2, 3 with a difference of one between each weight-value. Since the scoring plan of 3, 2, 1, 0 corresponds to the weight-values indicated by the test of internal consistency, it might be assumed that the scoring plan of 3, 2, 1, 0 would be the best method of scoring when the semi-external criteria tests of differentiation between sexes and among ages were applied. This assumption was not borne out, however, since it was shown in Table 9 that the 2, 1, 0, 0 scoring plan provided better differentiation between sexes and, with the exception of the 1, 0, 0, 0 plan, among ages.

Since the 3, 2, 1, 0 scoring plan did not prove to be the best method when tests involving semi-external criteria were applied, although it was shown to be the best method by a test involving internal consistency, it was considered possible that a separate continuum may exist at the opposite end of the response scale which measures some characteristic other than the cynicism defined in the validation of the test items. It is suggested that the opposite characteristic measured may be idealism. It is recommended that further study be conducted to determine the nature of the characteristic measured by this suggested continuum.

The assignment of weight-values of 2, 1, 0, 0 to the responses was selected as the scoring plan to be used for further investigation of the cynicism measured by the test for the following reasons:

1. As was shown in Table 9, the 2, 1, 0, 0 scoring plan differentiated to a much greater degree between sexes than other methods, and to only a slightly lesser degree than the 1, 0, 0, 0 plan among ages. An undetermined proportion of the 2, 1, 0, 0 F-value for sex was attributed to a tendency of the men to mark extremes, however.

2. The difference of one between the weight-values of 2, 1, and 0 corresponds to the difference indicated by the test of internal consistency based on the sigma scores of percentages of a normal distribution for the responses of this group of subjects, as was shown in Table 12.

3. The 2, 1, 0, 0 scoring plan provides a scale on which the degree of cynicism toward a testing item can be expressed by the subject but suggests that the expression of the degree of the opposite characteristic should be disregarded. When expression of the degree of the opposite characteristic is disregarded, the opposite characteristic can be considered on the same continuum with cynicism—the same difference existing between the two cynical responses as exists between the mildly cynical response and either of the responses of the opposite characteristic.

Recommendations

It is recommended that further investigation be conducted using external criteria rather than the semi-external criteria of sex and age considered in the present study. Cynical and idealistic groups of students might be selected from dormitories on the basis of ratings by fellow students.

Further analysis of the linearity of the continuum involved in the response scale could be undertaken by computing the total average score of all individuals responding to each of the four points provided on the scale and plotting these averages for noting linearity by inspection. The same procedure or more accurate mathematical treatment could be undertaken for each of the four scoring plans.

Summary

The responses of 400 college students to the 200 test items comprising the Fritz Test of Cynicism were analyzed to determine the optimal scoring plan for the test.

Four weight-value combinations for the four-point response scale were considered in selecting the best scoring plan. Assignment of weight-values of 2, 1, 0, 0 to the response scale—the highest value being assigned to the most cynical of the four responses—was designated as the first scoring plan; 3, 2, 1, 0 as the second; 4, 3, 1, 0 as the third; and 1, 0, 0, 0, 0 as the fourth. The 400 tests were scored and weighted by each of the plans.

Since the results of a preliminary study of the test of cynicism indicated the existence of highly significant differences between

sexes and among ages, the extent to which each scoring plan would differentiate between sexes and among ages was used as the initial criterion for evaluating the four scoring plans. The scores were classified by sex and age and treated by analysis of variance. It was found that the 2, 1, 0, 0 scoring plan differentiated to a much greater degree between sexes than the other methods, and to only a slightly lesser degree than the 1, 0, 0, 0 plan among ages. An undetermined proportion of the 2, 1, 0, 0 F-value for sex was attributed to a tendency of the men to mark extremes, however.

The coefficients of reliability computed for each of the four scoring plans from a sample of the scores of 50 subjects revealed no distinct scoring advantage concerning reliability for any one scoring plan.

When a normal distribution of responses for the group of subjects was assumed and the sigma scores of percentages were computed for each response, it was found that the weight-values based on sigma scores of percentages closely approximated an arithmetic sequence of 0, 1, 2, 3 with a difference of one between each weight-value.

Since the 3, 2, 1, 0 scoring plan did not prove to be the best method when tests involving semi-external criteria were applied, although it was shown to be the best method by a test involving internal consistency, it was considered possible that a separate continuum may exist at the opposite end of the response scale which measures some characteristic other than the cynicism defined in the validation of the test items, such as idealism.

The assignment of weight-values of 2, 1, 0, 0 to the response scale was tentatively selected as the scoring plan to be used for further investigation of cynicism.

IOWA STATE COLLEGE,
AMES, IOWA.

Literature Cited

1. Snedecor, G. W. 1946. Statistical Methods. 4th ed. p. 284 Ames, Iowa, Iowa State College Press.